

DEVICE FOR HIGHLY EFFICIENTLY ENCODING AUDIO SIGNAL

Patent number: JP2001142493

Publication date: 2001-05-25

Inventor: NAGAI KIYOTAKA; USAMI AKIRA; IIMURA KATSUHIKO

Applicant: MATSUSHITA ELECTRIC IND CO LTD

Classification:

- international: G10L19/00; G10L19/02; G10L11/00; H03M7/30

- european:

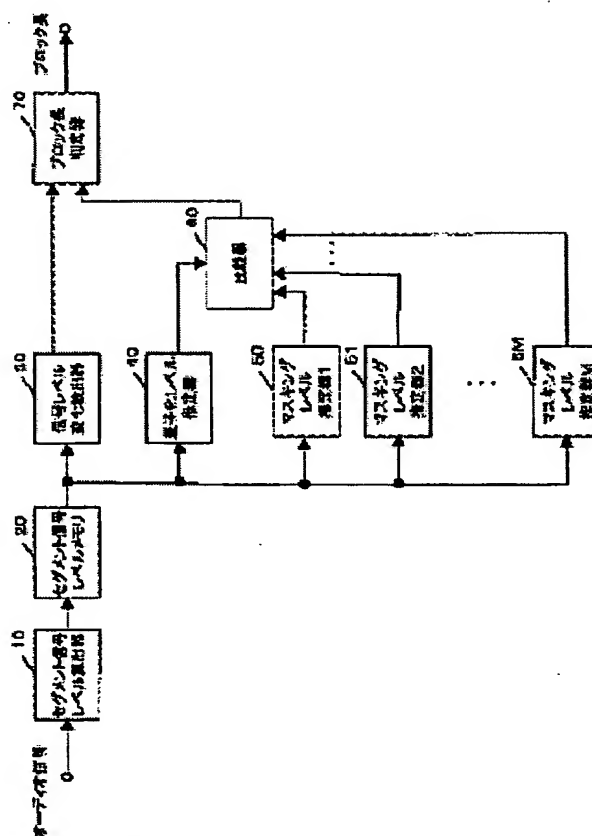
Application number: JP19990325093 19991116

Priority number(s):

Abstract of JP2001142493

PROBLEM TO BE SOLVED: To solve such problem that the tone quality is degraded by improperly selecting a short block length or missing a quick rise of a signal in a device which divides an audio signal into blocks to encode it.

SOLUTION: The audio signal is divided into segments, and signal levels of segments are calculated by a segment signal level calculator 10, and a quick rise of the signal is detected in a signal level change detector 30 by the change of signal levels of segments, and a quantization noise in segments preceding the segment where a quick rise of the signal has been detected is estimated by a quantization level estimator 40, and a masking level based on preceding segments is estimated in masking level estimators 50 and 51 to 5M by segments in which the quantization noise has been estimated, and the device is controlled by a block length decision device 70 so that the time of a block for encoding of the audio signal may be shortened in the case that the quantization noise is larger than the masking level.



Data supplied from the esp@cenet database - Patent Abstracts of Japan

【特許請求の範囲】

【請求項1】 オーディオ信号をブロックに分割して符号化する装置であって、

オーディオ信号を、あるいは高域通過フィルタもしくは帯域通過フィルタを通過したオーディオ信号をセグメントに分割し、前記セグメントの信号レベルを算出し、算出した前記信号レベルの変化から信号の急激な上昇を検出し、信号の急激な上昇が検出されたセグメントより前のセグメントにおける量子化ノイズを推定し、前記量子化ノイズを推定したセグメントにおける前記セグメントから前のセグメントによるマスキングレベルを推定し、前記量子化ノイズが前記マスキングレベルより大きい場合にオーディオ信号を符号化するブロックの時間を短くするように制御する手段を備えたことを特徴とするオーディオ信号高能率符号化装置。

【請求項2】 信号の上昇を検出するときには、高域通過フィルタあるいは帯域通過フィルタを通過したオーディオ信号をセグメントに分割した信号を用い、量子化ノイズ及びマスキングレベルを推定するときには、オーディオ信号をセグメントに分割した信号を用いることを特徴とする請求項1記載のオーディオ信号高能率符号化装置。

【請求項3】 信号の急激な上昇が検出されたセグメント i の j 個前のセグメントにおける量子化ノイズを $S(i-j) \times a$ (ただし、 $S(i)$ はセグメント i の信号レベル、 j は正整数で $(i-j)$ は最も長いブロックに対応するセグメントの範囲内の値、 a は所定の値) で推定し、前記 j 個前のセグメントにおける前記セグメントから k 個前のセグメント (ただし、 $0 \leq k \leq M$ 、 M は前向きマスキングを考慮するセグメント数) によるマスキングレベルを $S(i-j-k) \times b(k)$ (ただし、 $b(k)$ は k の減少関数で所定の値) で推定し、すべての k に対して前記量子化レベルが前記マスキングレベルより大きい場合にオーディオ信号を符号化するブロックの時間を短くするように制御する手段を備えたことを特徴とする請求項1または2記載のオーディオ信号高能率符号化装置。

【請求項4】 信号の急激な上昇が検出されたセグメント i の j 個前のセグメントにおける量子化ノイズを $S(i-j) \times a$ (ただし、 $S(i)$ はセグメント i の信号レベル、 j は正整数で $(i-j)$ は最も長いブロックに対応するセグメントの範囲内の値、 a は所定の値) で推定し、前記 j 個前のセグメントにおける前記セグメントから k 個前のセグメント (ただし、 $0 \leq k \leq M$ 、 M は前向きマスキングを考慮するセグメント数) によるマスキングレベルを $S(i-j-k) \times b(k)$ の和で (ただし、 $b(k)$ は k の減少関数で所定の値) 推定し、前記量子化レベルが前記マスキングレベルより大きい場合にオーディオ信号を符号化するブロックの時間を短くするように制御する手段を備えたことを特徴とする請求項

1または2記載のオーディオ信号高能率符号化装置。

【請求項5】 時間的にシフトした複数のセグメントの系列に対して、それぞれ上記制御を行う手段を備えたことを特徴とする請求項1ないし4のいずれかに記載のオーディオ信号の高能率符号化装置。

【請求項6】 セグメントの信号レベルは、セグメント内の信号の2乗値の和、絶対値の和、あるいは絶対値の最大であることを特徴とする請求項1ないし5のいずれかに記載のオーディオ信号高能率符号化装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、オーディオ信号をブロックに分割して高能率符号化するオーディオ信号高能率符号化装置に関する。

【0002】

【従来の技術】 近年、オーディオ信号の高能率符号化方式として変換符号化を利用した方式が広く用いられている。変換符号化を利用した方式としては、MPEG (Moving Picture Experts Group) 2のAAC (Advanced Audio Coding) やドルビーデジタルが挙げられる。また、サブバンドフィルタと変換符号化とを組み合わせたハイブリッド符号化方式を利用したものとしては、MPEG1およびMPEG2のレイヤ3とMD (ミニディスク) のATRACが挙げられる。

【0003】 図8に変換符号化を利用したオーディオ信号高能率符号化装置のブロック図を示す。図8において、301はブロック長決定部、302は周波数変換部、303は心理音響分析部、304は量子化及び符号化部である。以上のように構成されたオーディオ信号高能率符号化装置について、その動作を以下に述べる。

【0004】 入力された時間軸上のオーディオ信号は、周波数変換部302においてブロック長決定部301で決定された長さのブロックに分割され、直交変換されて周波数軸のスペクトル係数に変換される。ハイブリッド符号化方式の場合には、入力されたオーディオ信号はフィルタバンクでサブバンド信号に分割され、各サブバンド信号はブロック長決定部301で決定された長さのブロックに分割され、直交変換されて周波数軸のスペクトル係数に変換される。直交変換としては、主としてMDCT (変形離散コサイン変換) やFFT (高速フーリエ変換) が用いられる。ブロック長決定部301では、信号の変化に応じてブロック長、すなわち直交変換長を変化させ、プリエコーと呼ばれる疑似信号による音質の劣化を防止する。

【0005】 図9は、ブロック長とプリエコーの関係を示す図である。同図は直交変換としてMDCTをもちいたときの図で、隣接ブロックは50%オーバーラップしている。同図(a)は入力信号波形を、(b)は(a)の入力信号を長いブロック長を使って符号化・復号化した信号波形と変換に用いた窓の波形を、(c)は短いブ

ロック長を使って符号化・復号化した信号波形と交換に用いた窓の波形を示す。また、同図(b)でTLは長いブロック長の時間を、同図(c)でTSは短いブロック長の時間を示す。図9ではTSはTLの1/4の時間である。

【0006】図9(a)に示すような急激な立ち上がりを含む信号を、同図(b)のような長いブロック長で交換符号化すると、振幅の大きな部分の引き起こす量子化ノイズが振幅の小さい部分に広がり疑似信号を発生する。一方、同図(c)に示すように短いブロック長で交換符号化すると、振幅の大きな部分の引き起こす量子化ノイズは短いブロックの中に閉じ込められる。大きな振幅の信号による量子化ノイズは交換ブロック全体に発生するが、前向きマスキング(フォワードマスキング)の方が後向きマスキング(バックワードマスキング)より作用を及ぼす時間が長いので大きな信号の後のノイズは大きな信号の前に発生するノイズと比較して知覚されにくい。大きな信号の前に発生するノイズは本来の信号が聞こえる前に聞こえるのでプリエコーと呼ばれ、品質を大きく劣化させる。

【0007】したがって、交換符号化方式やハイブリッド符号化方式では、急激な立ち上がり信号に対しては短いブロック長を選択することによりプリエコーを抑圧する。しかしながら、定常的な信号に対して短いブロック長を適用すると符号化効率が低下し、S/N(signal to noise ratio)が劣化する。なぜなら、符号化データとしては、量子化されたスペクトル情報の外に補助情報が1ブロックに1組必要なので、短いブロック長を選択したときは符号化データの中で補助情報の占める割合が増加し、符号化効率が低下するからである。

【0008】従来のブロック長決定部としては、例えば特開平3-263926号公報に記載されたものが知られている。

【0009】図10に従来のブロック長決定部のブロック図を示す。図10において、401はセグメント信号レベル算出器、402はセグメント信号レベルメモリ、403は信号レベル変化検出器である。以上のように構成された従来のブロック長決定部について、その動作を以下に述べる。

【0010】入力されたオーディオ信号は、セグメント信号レベル算出器401で最も短いブロックより小さい時間のセグメントに分割され、各セグメントの信号レベルをセグメント内の信号の2乗値の和、すなわちエネルギーで算出する。セグメント信号レベルメモリ402ではセグメント信号レベル算出器401で算出された各セグメントの信号レベルをメモリに記憶する。信号レベル変化検出器403では、セグメント信号レベルメモリ402から読み出したセグメントの信号レベルを用いて、隣接するセグメントの信号レベルの比を求め、この値が閾値を越えるときには短いブロック長を表す信号を出力

し、そうでないときには長いブロック長を表す信号を出力する。以上のようにして信号の急激な上昇を検出したときには短いブロック長を表す信号を出力することにより、プリエコーを抑圧することができる。

【0011】図8の心理音響分析部303では、オーディオ信号、ブロック長、スペクトル係数を入力として、心理音響モデルに基づいてスペクトル係数のマスキングレベル、すなわち許容量子化ノイズレベルを算出し、出力する。

【0012】量子化及び符号化部304では、心理音響分析部303の出力である許容量子化ノイズレベルに基づいてスペクトル係数の量子化を行い、ハフマン符号化等の符号化処理を行い、高能率符号化データを出力する。

【0013】

【発明が解決しようとする課題】しかしながら、上記従来のオーディオ信号高能率符号化装置では、音声の母音部のように定常的ではあるが、信号レベルが急激に変化する信号に対しては、短いブロック長を選択する場合があり、符号化効率の高い長いブロック長を選択した場合と比較して、音質(S/N)が劣化する、という課題があった。

【0014】また、ブロック長を決定する時に用いるセグメントと信号の相対時間関係によって信号の急激な変化を見逃す場合があり、これによってプリエコーが発生し、音質が劣化するという課題があった。

【0015】本発明は上記問題点を解決するもので、短いブロック長の不適切な選択による符号化効率の低下による音質劣化を低減した高品質なオーディオ信号高能率符号化装置を提供することを目的とする。

【0016】また、信号の急激な立ち上がりを見逃すことにより引き起こされるプリエコーによる音質劣化を低減したオーディオ信号高能率符号化装置を提供することを目的とする。

【0017】

【課題を解決するための手段】この課題を解決するために、本発明のオーディオ信号高能率符号化装置は、オーディオ信号をブロックに分割して符号化する装置であって、オーディオ信号を、あるいは高域通過フィルタもしくは帯域通過フィルタを通過したオーディオ信号をセグメントに分割し、セグメントの信号レベルを算出し、算出した前記信号レベルの変化から信号の急激な上昇を検出し、信号の急激な上昇が検出されたセグメントより前のセグメントにおける量子化ノイズを推定し、前記量子化ノイズを推定したセグメントにおける前記セグメントから前のセグメントによるマスキングレベルを推定し、前記量子化ノイズが前記マスキングレベルより大きい場合にオーディオ信号を符号化するブロックの時間を短くするように制御する手段を備えたものである。

【0018】また、本発明のオーディオ信号高能率符号

化装置は、信号の上昇を検出するときには、高域通過フィルタあるいは帯域通過フィルタを通過したオーディオ信号をセグメントに分割した信号をもちい、量子化ノイズ及びマスキングレベルを推定するときには、オーディオ信号をセグメントに分割した信号をもちいるものである。

【0019】また、本発明のオーディオ信号高能率符号化装置は、量子化ノイズを推定したセグメントから前のセグメントによるマスキングレベルを直接、あるいは累積して前記量子化ノイズと比較する手段を備えたものである。

【0020】また、本発明のオーディオ信号高能率符号化装置は、時間的にシフトした複数のセグメントの系列に対して、それぞれ上記制御を行う手段を備えたものである。

【0021】以上の発明において、短いブロック長の適切な選択を可能とし、また、信号の立ち上がりの変化を精度良く検出するので、高音質なオーディオ信号高能率符号化装置を実現することができる。

【0022】

【発明の実施の形態】以下、本発明の実施の形態について、図面を用いて説明する。

【0023】本発明のオーディオ信号高能率符号化装置の全体の構成は、図8に示す一般的な構成と同一である。

【0024】図8において入力された時間軸のオーディオ信号は、周波数変換部302においてブロック長決定部301で決定された長さのブロックで直交変換されて、周波数軸のスペクトル係数に変換される。

【0025】本実施の形態では、直交変換としてMDC Tを用いる。ブロック長決定部301では、信号の変化に応じてブロック長すなわち直交変換長を変化させ、ブリエコーと呼ばれる疑似信号の発生による音質の劣化を防止する。

【0026】なお、ハイブリッド符号化方式の場合には、入力されたオーディオ信号はフィルタバンク（図示せず）でサブバンド信号に分割され、各サブバンド信号はブロック長決定部301で決定された長さのブロックに分割され、周波数変換部302で直交変換されてスペクトル係数に変換される。

【0027】心理音響分析部303では、オーディオ信号、ブロック長、スペクトル係数を入力として、心理音響モデルに基づいてスペクトル係数のマスキングレベル、すなわち許容量子化ノイズレベルを算出し、出力する。

【0028】量子化及び符号化部304では心理音響分析部303の出力である許容量子化ノイズレベルに基づいてスペクトル係数の量子化を行い、ハフマン符号化等の符号化処理を行い、高能率符号化データを出力する。

【0029】以下、本発明の特徴であるブロック長決定

部301について、図面を用いて詳細に説明する。

【0030】（実施の形態1）図1は、本発明の実施の形態1によるオーディオ信号高能率符号化装置のブロック長決定部の構成を示すブロック図である。

【0031】図1において、10はセグメント信号レベル算出器、20はセグメント信号レベルメモリ、30は信号レベル変化検出器、40は量子化レベル推定器、50、51～5Mはマスキングレベル推定器、60は比較器、70はブロック長判定器である。

【0032】以上のように構成された本実施の形態におけるブロック長決定部について、以下その動作を述べる。

【0033】入力されたオーディオ信号は、セグメント信号レベル算出器10で最も短いブロックの時間と同じかそれより小さい時間のセグメントに分割され、各セグメントの信号レベルをセグメント内の信号の2乗値の和で算出される。

【0034】図2に本実施の形態におけるブロックとセグメントの関係を示す。同図で上から順に長いブロック、短いブロック、セグメントを表し、破線は相対時間関係を示す。本実施の形態では、直交変換としてMDC Tを用いているため、各ブロックは50%オーバーラップしている。長いブロックは、通常512サンプルから2048サンプルで、また短いブロックは通常64サンプルから256サンプルで構成される。本実施の形態では、長いブロックは1024サンプル、短いブロックは256サンプル、セグメントは128サンプルで構成されている。すなわち、短いブロック長（TS）は長いブロック長（TL）の1/4の時間で、セグメント長（T）は短いブロック長（TS）の1/2の時間である。

【0035】なお、セグメント信号レベル算出器10で算出されるセグメントの信号レベルとしては、ダイナミックレンジを小さくするため、あるいは処理量を減らすために、上記セグメント内の信号の2乗値の和の代わりに信号の絶対値の和、あるいは信号の絶対値の最大を用いてもよい。また、ハイブリッド符号化方式の場合には、サブバンドフィルタ（帯域通過フィルタ）の出力であるサブバンド信号をセグメント信号レベル算出器10に入力する。

【0036】次に、セグメント信号レベルメモリ20では、セグメント信号レベル算出器10で算出された各セグメントの信号レベルをメモリに記憶する。ここで、セグメントi（iは任意の整数）の信号レベルをS（i）とする。信号レベル検出器30では、セグメント信号レベルメモリ20から読み出した信号レベルを用いて、セグメントiの信号レベルS（i）に対する直前のセグメントの信号レベルS（i-1）の比、すなわちS（i）/S（i-1）を求め、この値が所定の閾値を越えるときには、急激な信号の上昇が検出されたことを表す信号

をオンにして出力する。

【0037】量子化レベル推定器40では、信号の急激な上昇を検出したセグメント*i*の直前のセグメントにおける量子化ノイズを $S(i) \times a$ （ただし、*a*は所定の値）で推定し、量子化ノイズを出力する。

【0038】マスキングレベル推定器50では、セグメント*i*の直前のセグメントの信号レベル $S(i-1)$ によるセグメント*i*におけるマスキングレベルを $S(i-1) \times b(0)$ で推定し、マスキングレベルを出力する。同様に、マスキングレベル推定器5*k*では、セグメント*i*の直前のセグメントから*k*個前（ただし、 $0 \leq k \leq M$ で、*M*は前向きマスキングを考慮するセグメント数）のセグメントの信号レベル $S(i-1-k)$ によるセグメント*i*の直前のセグメントにおける前向きマスキングのマスキングレベルを $S(i-1-k) \times b(k)$ で推定し、マスキングレベルを出力する。ここで*b(k)*は*k*の減少関数で所定の値である。

【0039】比較器60では、量子化レベル推定器40からの量子化ノイズとマスキングレベル推定器50ないし5*M*からのマスキングレベルとを比較し、 $0 \leq k \leq M$ のすべての*k*に対して量子化ノイズの方がマスキングレベルより大きいときに、量子化ノイズがマスキングレベルより大きいことを表す信号をオンにして出力する。

【0040】ブロック長判定器70では、信号レベル検出器30からの信号の急激な上昇を表す信号がオンで、かつ比較器60からの量子化ノイズがマスキングレベルより大きいことを表す信号がオンのときに短いブロック長を表す信号を出力し、そうでないときには長いブロック長を表す信号を出力する。

【0041】図3は、実施の形態1の動作を説明する図である。同図(a)はセグメント信号レベルメモリ20に記憶されたセグメント*i*の信号レベル $S(i)$ を示す。同図(a)に示すように、 $S(i)/S(i-1)$ は所定の閾値より大きいので信号レベル変化検出器30の出力信号はオンとなる。同図(b)に信号レベルの急激な上昇が検出されたセグメント*i*の信号レベル $S(i)$ が直前のセグメント(*i-1*)に及ぼす量子化ノイズ $S(i) \times a$ を短い破線で示す。また、前記直前のセグメントの信号レベル $S(i-1)$ によるマスキングレベル $S(i-1) \times b(0)$ を長い破線で示す。同図(b)に示すように量子化ノイズの方が前記マスキングレベルより大きい。同図(c)に前記直前のセグメントの1つ前のセグメントの信号レベル $S(i-2)$ が前記直前のセグメントに及ぼすマスキングレベル $S(i-2) \times b(1)$ を一点鎖線で示す。同図(c)に示すように量子化ノイズの方が前記マスキングレベルより大きい。同図(d)に前記直前のセグメントの2つ前のセグメントの信号レベル $S(i-3)$ が前記直前のセグメントに及ぼすマスキングレベル $S(i-3) \times b(2)$ を二点鎖線で示す。この場合、同図(d)に示すように前

記マスキングレベルの方が量子化ノイズより大きい。

【0042】図3に示す例の場合、マスキングレベルの方が量子化ノイズより大きいので信号の急激な立ち上がりの直前の量子化ノイズは聞こえない。したがって、本実施の形態ではブロック長判定器70長いブロック長を表す信号を出力する。これに対して、従来例では前向きマスキングを考慮していないので短いブロック長を表す信号を出力し、符号化効率を低下させていた。

【0043】なお、実施の形態1では信号の急激な上昇が検出されたセグメントの直前セグメントにおける量子化ノイズとマスキングレベルを推定したが、信号の急激な上昇が検出されたセグメントの*j*個前のセグメント（ただし、*j*は正整数で最も長いブロックに対応するセグメントの範囲内の値）に対して量子化ノイズとマスキングレベルを推定してもよい。

【0044】以上のように本実施の形態では、量子化レベル推定器40、マスキングレベル推定器50～5*M*、比較器60を設けることにより、信号レベルの高い部分が低い部分に引き起こす量子化ノイズと前向きマスキングによるマスキングレベルとを比較することにより、短いブロック長の不適切な選択による符号化効率の低下による音質劣化を低減することができる。

【0045】すなわち、本実施の形態によれば、信号レベルの急激な変化に加えて、前向きマスキングを考慮してブロック長を決定することにより、信号レベルの急激な変化を検出しても、信号レベルの高い部分が低い部分に引き起こす量子化ノイズが前向きマスキングによってマスキングされる場合には、符号化効率の高い長いブロック長を選択することにより、音声、特に男性音の音質(*S/N*)を向上させることができる。

【0046】（実施の形態2）図4は、本発明の実施の形態2によるオーディオ信号高能率符号化装置のブロック長決定部の構成を示すブロック図である。

【0047】図4において、10と90はセグメント信号レベル算出器、20と95はセグメント信号レベルメモリ、30は信号レベル変化検出器、40は量子化レベル推定器、50、51～5*M*はマスキングレベル推定器、60は比較器、70はブロック長判定器、80は高域通過フィルタ、85は遅延メモリである。

【0048】図1に示した実施の形態1と図4に示した実施の形態2の構成の違いは、実施の形態1に対して、実施の形態2では、高域通過フィルタ80、遅延メモリ85、セグメント信号レベル算出器90、セグメント信号レベルメモリ95が追加されている点である。

【0049】以上のように構成された本実施の形態のブロック長決定部について、以下その動作を述べる。

【0050】入力されたオーディオ信号は、高域通過フィルタ80で低域成分を除去される。なお、高域通過フィルタ80の代わりに帯域通過フィルタを用いてもよ

【0051】次にセグメント信号レベル算出器10で高域通過フィルタ80を通過したオーディオ信号を最も短いブロックの時間と同じかそれより小さい時間のセグメントに分割し、各セグメントの信号レベルをセグメント内の信号の2乗値の和で算出する。セグメント信号レベルメモリ20、信号レベル検出器30の動作は実施の形態1と同一であるので説明を省略する。

【0052】一般に信号の急激な変化は、低域周波数よりも高域周波数に特徴的に現れるので、低域成分を除去した信号を用いて変化の検出を行うことにより、検出の精度を高めることができる。

【0053】一方、遅延メモリ85では入力されたオーディオ信号を高域通過フィルタ80による遅延時間と同じ時間だけ遅延させた信号を出力する。

【0054】次にセグメント信号レベル算出器90で遅延メモリ85の出力信号を最も短いブロックの時間と同じかそれより小さい時間のセグメントに分割し、各セグメントの信号レベルをセグメント内の信号の2乗値の和で算出する。

【0055】なお、セグメント信号レベル算出器10、あるいは90で算出されるセグメントの信号レベルとしては、ダイナミックレンジを小さくするため、あるいは処理量を減らすために、上記セグメント内の信号の2乗値の和の代わりに信号の絶対値の和、あるいは信号の絶対値の最大を用いてもよい。

【0056】遅延メモリ85を挿入しているのでセグメント信号レベル算出器10で分割されるセグメントと、セグメント信号レベル算出器90で分割されるセグメントのセグメント分割タイミングは同一で、実施の形態1と同様に、図2で示される。

【0057】また、本実施の形態の周波数変換部302（図8）では入力されたオーディオ信号を高域通過フィルタによる遅延時間分だけ遅らす遅延メモリ（図示せず）を通した後で、MDCTを行うことにより、MDCTの変換ブロックとセグメントの時間関係を図2の関係に保つ。

【0058】次にセグメント信号レベルメモリ95では、セグメント信号レベル算出器90で算出された各セグメントの信号レベルをメモリに記憶する。

【0059】量子化レベル推定器40、マスキングレベル推定器50、51～5M、比較器60、ブロック長判定器70の動作は実施の形態1と同一であるので説明を省略する。

【0060】以上のように実施の形態2では、信号の上昇を検出するときには高域通過フィルタ80を通過したオーディオ信号をセグメントに分割した信号を用いることにより、実施の形態1と比較して信号の変化の検出精度を高めることができる。

【0061】以上のように実施の形態2によれば、精度の高い信号レベルの急激な変化の検出に加えて、前向き

マスキングを考慮してブロック長を決定することにより、信号レベルの急激な変化を検出しても、信号レベルの高い部分が低い部分に引き起こす量子化ノイズが前向きマスキングによってマスキングされる場合には、符号化効率の高い長いブロック長を選択することにより、音声、特に男性音の音質（S/N）を向上させることができる。

【0062】（実施の形態3）図5は、本発明の実施の形態3によるオーディオ信号高能率符号化装置のブロック長決定部の構成を示すブロック図である。

【0063】図5において、10はセグメント信号レベル算出器、20はセグメント信号レベルメモリ、30は信号レベル変化検出器、40は量子化レベル推定器、50、51～5Mはマスキングレベル推定器、60は比較器、70はブロック長判定器、100は加算器である。

【0064】図1に示す実施の形態1と図5に示す実施の形態3の構成の違いは、実施の形態3では、マスキングレベル推定器50ないし5Mの出力に加算器100が追加されている点である。

【0065】以上のように構成された本発明の実施の形態3を示すブロック長決定部について、以下その動作を述べる。

【0066】入力されたオーディオ信号は、セグメント信号レベル算出器10で最も短いブロックの時間と同じかそれより小さい時間のセグメントに分割され、各セグメントの信号レベルをセグメント内の信号の2乗値の和で算出される。

【0067】図2に実施の形態3におけるブロックとセグメントの関係を示す。この関係は実施の形態1と同一である。実施の形態3では、直交変換としてMDCTをもちい、長いブロックは1024サンプル、短いブロックは256サンプル、セグメントは128サンプルで構成されている。

【0068】なお、セグメント信号レベル算出器10で算出されるセグメントの信号レベルとしては、ダイナミックレンジを小さくするため、あるいは処理量を減らすために、上記セグメント内の信号の2乗値の和に代わりに信号の絶対値の和、あるいは信号の絶対値の最大を用いてもよい。また、ハイブリッド符号化方式の場合には、サブバンドフィルタ（帯域通過フィルタ）の出力であるサブバンド信号をセグメント信号レベル算出器10に入力する。

【0069】次にセグメント信号レベルメモリ20ではセグメント信号レベル算出器10で算出された各セグメントの信号レベルをメモリに記憶する。ここで、セグメントiの信号レベルを $S(i)$ とする。信号レベル検出器30では、セグメント信号レベルメモリ20から読み出した信号レベルを用いて、セグメントiの信号レベル $S(i)$ に対する直前のセグメントの信号レベル $S(i-1)$ の比、すなわち $S(i)/S(i-1)$ を求め、

この値が閾値を越えるときには、急激な信号の上昇が検出されたことを表す信号をオンにして出力する。

【0070】量子化レベル推定器40では、信号の急激な上昇を検出したセグメント*i*の直前のセグメントにおける量子化ノイズを $S(i) \times a$ （ただし、*a*は所定の値）で推定し、前記量子化ノイズを出力する。

【0071】マスキングレベル推定器50では、前記直前のセグメントの信号レベル $S(i-1)$ による前記セグメントにおけるマスキングレベルを $S(i-1) \times b(0)$ で推定し、前記マスキングレベルを出力する。同様に、マスキングレベル推定器5*k*では前記直前のセグメントより*k*個前（ただし、 $0 \leq k \leq M$ で、*M*は前向きマスキングを考慮するセグメント数）のセグメントの信号レベル $S(i-1-k)$ による前記直前のセグメントにおける前向きマスキングのマスキングレベルを $S(i-1-k) \times b(k)$ で推定し、前記マスキングレベルを出力する。ここで*b(k)*は*k*の減少関数で所定の値である。

【0072】加算器100では、 $(M+1)$ 個のマスキングレベル推定器50ないし5*M*の出力であるマスキングレベルの和（累積値）を算出し、出力する。

【0073】比較器60では、量子化レベル推定器40からの量子化ノイズと加算器100からのマスキングレベルの累積値とを比較し、量子化ノイズの方がマスキングレベルの累積値より大きいときに量子化ノイズがマスキングレベルより大きいことを表す信号をオンにして出力する。

【0074】ブロック長判定器70では、信号レベル検出器30からの信号の急激な上昇を表す信号がオンで、かつ比較器60からの量子化ノイズがマスキングレベルより大きいことを表す信号がオンのときに短いブロック長を表す信号を出力し、そうでないときには長いブロック長を表す信号を出力する。

【0075】以上のように実施の形態3のブロック長決定部は、信号の急激な立ち上がりが検出されてもマスキングレベルの方が量子化ノイズより大きいときには前記量子化ノイズは聞こえないので符号化効率の高い長いブロック長を選択し、出力する。

【0076】すなわち、実施の形態3では、量子化レベル推定器40、マスキングレベル推定器50～5*M*、比較器60、加算器100を設けることにより、信号レベルの高い部分が低い部分に引き起こす量子化ノイズと前向きマスキングによるマスキングレベルとを比較することにより、短いブロック長の不適切な選択による符号化効率の低下による音質劣化を低減することができる。

【0077】なお、実施の形態3では信号の急激な上昇が検出されたセグメントの直前セグメントにおける量子化ノイズとマスキングレベルを推定したが、信号の急激な上昇が検出されたセグメントの*j*個前のセグメント

（ただし、*j*は正整数で最も長いブロックに対応するセ

グメントの範囲内の値）に対して量子化ノイズとマスキングレベルを推定してもよい。

【0078】実施の形態3では、実施の形態1と比較して加算器100を追加しなければならないが、加算器100でマスキングレベルの累積値を算出することにより、精度の高いマスキングレベルの算出を行うことができる。

【0079】以上のように本実施の形態によれば、信号レベルの急激な変化に加えて、前向きマスキングを考慮してブロック長を決定することにより、信号レベルの急激な変化を検出しても、信号レベルの高い部分が低い部分に引き起こす量子化ノイズがマスキングされる場合には、符号化効率の高い長いブロック長を選択することにより、音声、特に男性音の音質（*S/N*）を向上させることができる。

【0080】（実施の形態4）図6は、本発明の実施の形態4によるオーディオ信号高能率符号化装置のブロック長決定部の構成を示すブロック図である。

【0081】図6において、10と110はセグメント信号レベル算出器、20と120はセグメント信号レベルメモリ、30と130は信号レベル変化検出器、40と140は量子化レベル推定器、50～5*M*と150～15*M*はマスキングレベル推定器、60と160は比較器、70と170はブロック長判定器、100と200は加算器、210は統合判定器、220は第1のセグメント系列によるブロック長算出部、230は第2のセグメント系列によるブロック長算出部である。

【0082】第1のセグメント系列によるブロック長判定部220は、セグメント信号レベル算出器10、セグメント信号レベルメモリ20、信号レベル変化検出器30、量子化レベル推定器40、マスキングレベル推定器50～5*M*、比較器60、ブロック長判定器70、加算器100から構成されている。また、第2のセグメント系列によるブロック長判定部230は、セグメント信号レベル算出器110、セグメント信号レベルメモリ120、信号レベル変化検出器130、量子化レベル推定器140、マスキングレベル推定器150～15*M*、比較器160、ブロック長判定器170、加算器200から構成されている。

【0083】以上のように構成された本発明の実施の形態4を示すブロック長決定部について、以下その動作を述べる。

【0084】図5に示す実施の形態3と図6に示す実施の形態4の構成の違いは、実施の形態4では実施の形態3の構成（図6では第1のセグメント系列によるブロック長判定部220に相当）に判定統合器210と第2のセグメント系列によるブロック長算出部230が追加されている点である。

【0085】第1のセグメント系列によるブロック長算出部220の動作は、実施の形態3と同一であるので説

明を省略する。

【0086】また、第1のセグメント系列によるブロック長算出部220と第2のセグメント系列によるブロック長算出部230の構成は同一で、両者の違いは、オーディオ信号をセグメントに分割するときの分割の仕方が異なる点である。すなわち、セグメント信号レベル算出器10とセグメント信号レベル算出器110の動作が異なる点である。

【0087】第2のセグメント系列によるブロック長算出部230のセグメントは、第1のセグメント系列によるブロック長算出部220のセグメントを時間的にシフトしたものである。ここでは、セグメント信号レベル算出器10で分割されるセグメントを第1のセグメント系列、また、セグメント信号レベル算出器110で分割される前記第1のセグメント系列を時間的にシフトしたセグメント系列を第2のセグメント系列と呼ぶ。本実施の形態では、前記シフト量はセグメントの時間の1/2である。第2のセグメント系列によるブロック長算出部230のセグメント信号レベル算出器110における上記セグメントに分割する動作を除いた動作は、第1のセグメント系列によるブロック長算出部220と同一であるので説明を省略する。

【0088】図7は実施の形態4のブロック長決定部の動作を説明する図である。同図(a)は入力されたオーディオ信号の振幅を示す。同図(b)はセグメント信号レベル算出器10で算出された第1のセグメント系列のセグメントiの信号レベル $S1(i)$ を表す。また、同図(c)はセグメント信号レベル算出器110で算出された第2のセグメント系列のセグメントjの信号レベル $S2(j)$ を表す。同図(b)に示すように第1のセグメント系列によるブロック長算出部220では、 $S1(i)/S1(i-1)$ 及び $S1(i+1)/S1(i)$ の値が所定の閾値以下なので、信号レベル変化検出器30では急激な信号レベルの上昇を検出することができない。

【0089】これに対して同図(c)に示すように第2のセグメント系列によるブロック長算出部230では $S2(j)/S2(j-1)$ が所定の閾値より大きいので信号レベル変化検出器110で信号レベルの急激な上昇を検出することができる。判定統合器210ではブロック長判定器70とブロック長判定器170の出力を統合し、結果を出力する。本実施の形態では、判定統合器210はブロック長判定器70とブロック長判定器170の一方が短いブロック長を表す信号を出力したときに短いブロック長を表す信号を出力する。

【0090】以上のように実施の形態4では、ブロック長を判定するセグメントを第1のセグメントに対して時間的にシフトした第2のセグメントに対しても行うことにより、信号とセグメントの相対時間関係による信号レベルの急激な変化の見逃しを防止し、ブリエコーによる

音質劣化の少ない高音質なオーディオ信号符号化装置を実現することができる。

【0091】なお、以上の説明では、セグメントの時間(T)の1/2の時間シフトする例を示したが、セグメントをシフトする時間をそれ以外の値、例えばセグメントの時間の1/4シフトし、4個のセグメント系列に対してブロック長判定処理を行い、その結果を判定統合器で統合してもよい。

【0092】

【発明の効果】以上のように本発明によれば、信号レベルの変化に加えて、前向きマスキングを考慮してブロック長を決定することにより、信号レベルの急激な上昇を検出しても、信号レベルの高い部分が低い部分に引き起こす量子化ノイズがマスキングされる場合には、符号化効率の高い長いブロック長を選択することにより、音声、特に男性音の音質(S/N)を向上させることができる。

【0093】また、ブロック長を判定する処理を元のセグメントを時間的にシフトしたセグメントに対しても行うことにより、信号とセグメントの相対時間関係による信号レベルの急激な変化の見逃しを防止し、ブリエコーによる音質劣化のない高品質なオーディオ信号高能率符号化装置を実現できる。

【図面の簡単な説明】

【図1】本発明の実施の形態1によるオーディオ信号高能率符号化装置のブロック長決定部の構成を示すブロック図

【図2】同オーディオ信号高能率符号化装置における長いブロック、短いブロック、セグメントの関係を示す図

【図3】同オーディオ信号高能率符号化装置のブロック長決定部の動作を説明する図

【図4】本発明の実施の形態2によるオーディオ信号高能率符号化装置のブロック長決定部の構成を示すブロック図

【図5】同実施の形態3によるオーディオ信号高能率符号化装置のブロック長決定部の構成を示すブロック図

【図6】同実施の形態4によるオーディオ信号高能率符号化装置のブロック長決定部の構成を示すブロック図

【図7】同ブロック長決定部の動作を説明する図

【図8】本発明におけるオーディオ信号高能率符号化装置の構成を示すブロック図

【図9】ブロック長とブリエコーの関係を示す図

【図10】従来のブロック長決定部の構成を示す図

【符号の説明】

10, 90, 110 セグメント信号レベル算出器

20, 95, 120 セグメント信号レベルメモリ

30, 130 信号レベル変化検出器

40, 140 量子化レベル推定器

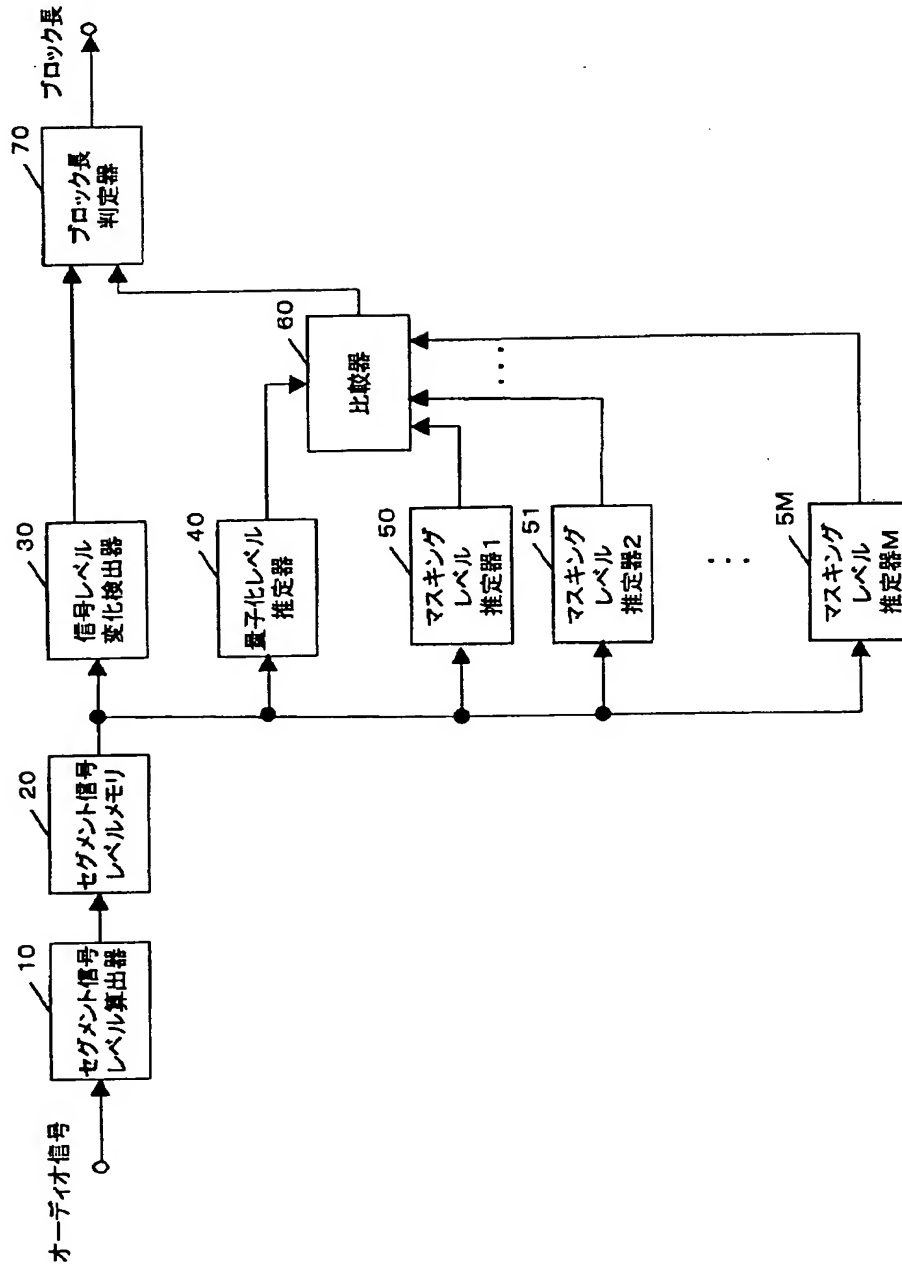
50, 51~5M, 150~15M マスキングレベル

推定器

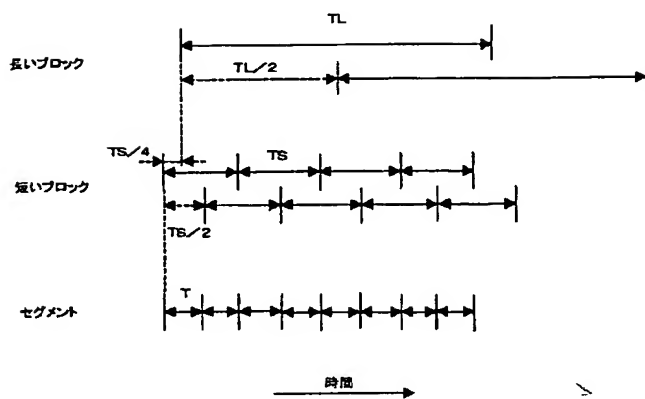
60, 160 比較器
 70, 170 ブロック長判定器
 80 高域通過フィルタ
 85 遅延メモリ
 100, 200 加算器
 210 判定統合器

* 220 第1のセグメント系列によるブロック長算出部
 230 第2のセグメント系列によるブロック長算出部
 301 ブロック長決定部
 302 周波数変換部
 303 心理音響分析部
 * 304 量子化及び符号化部

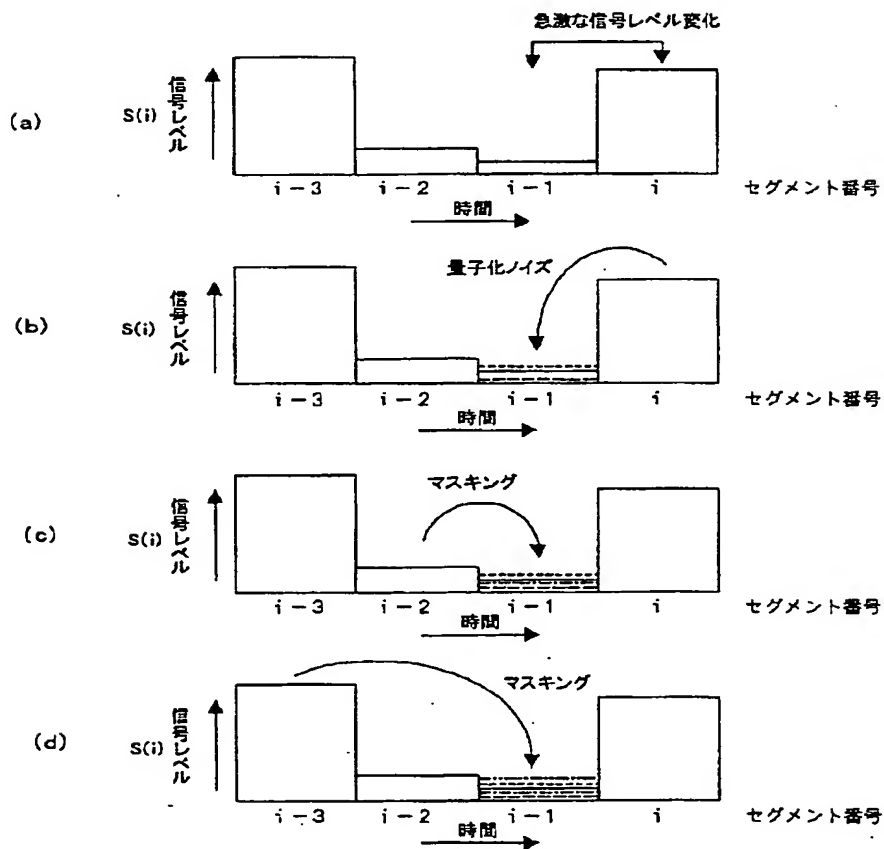
【図1】



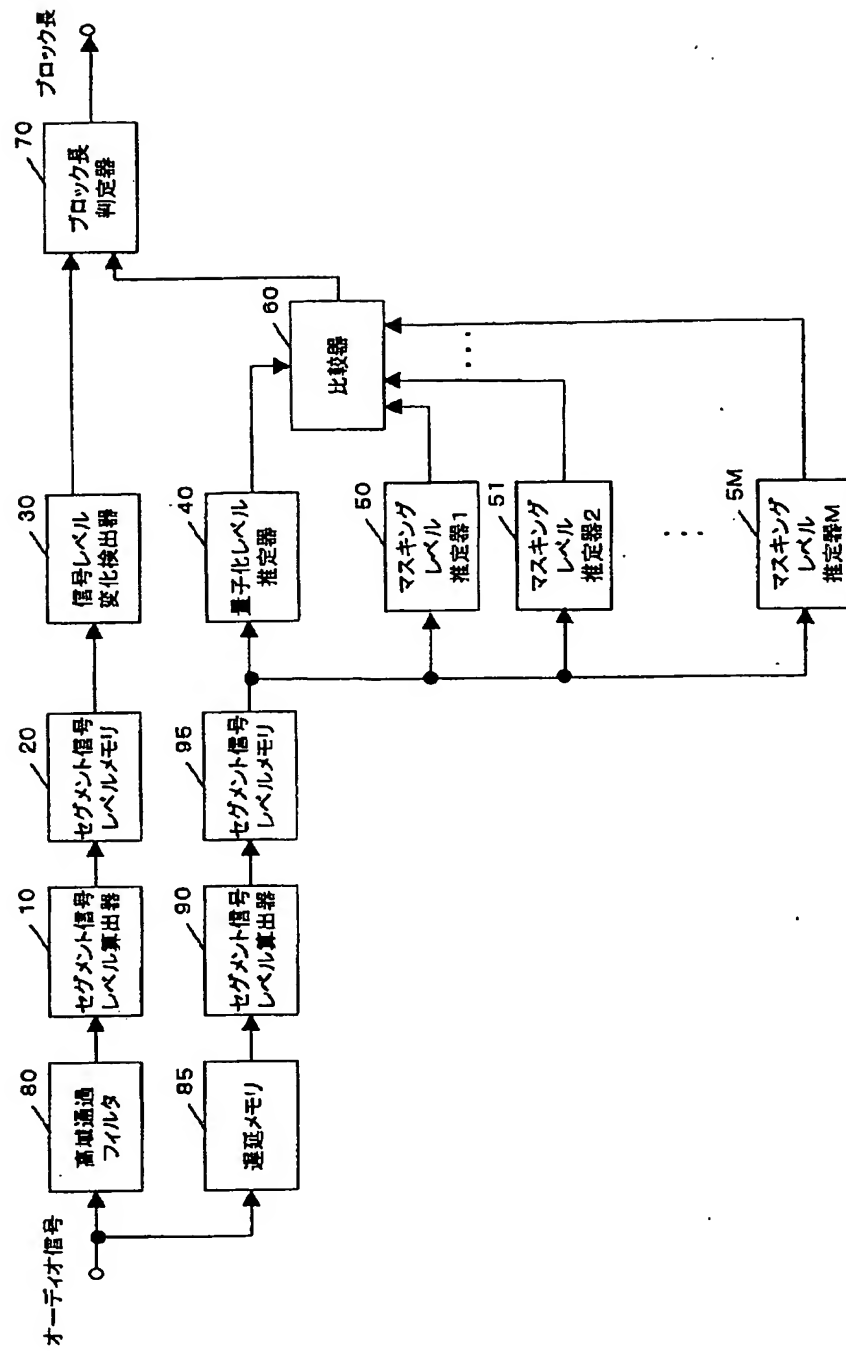
【図2】



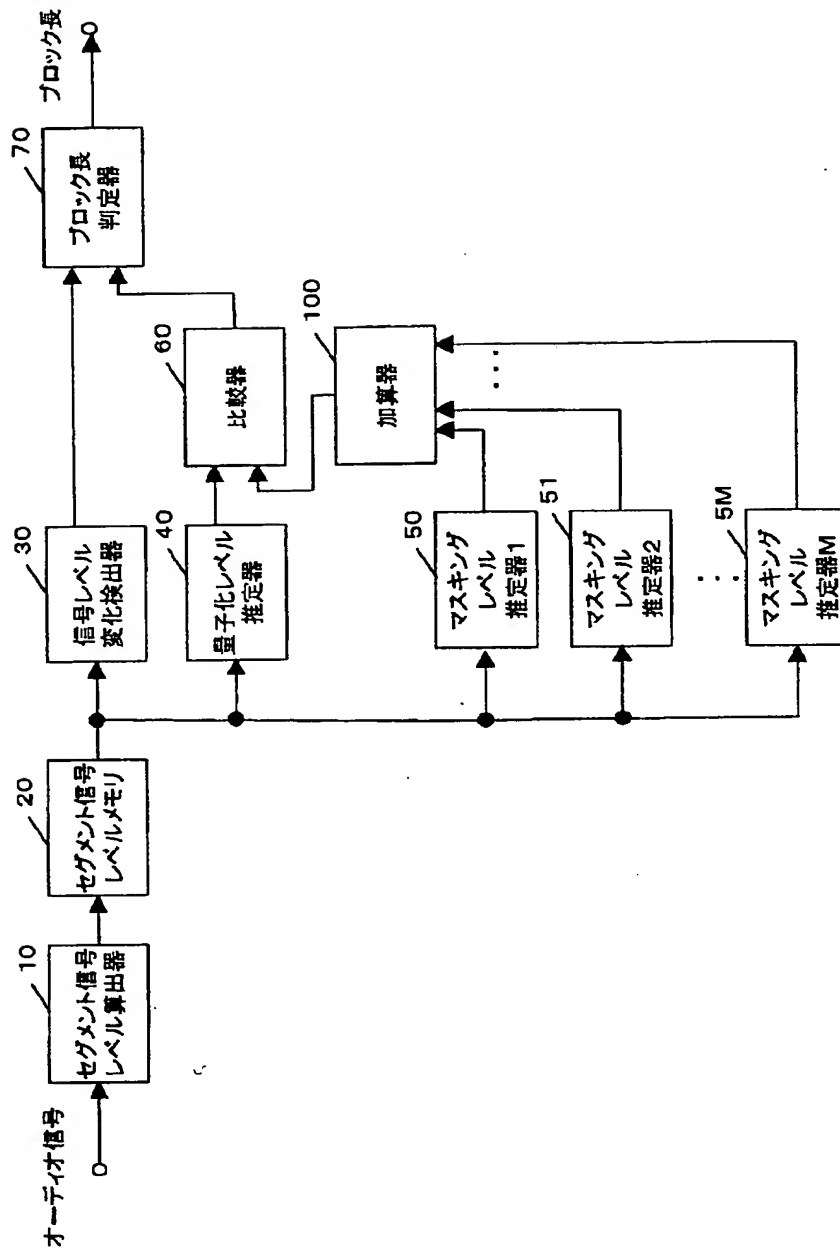
【図3】



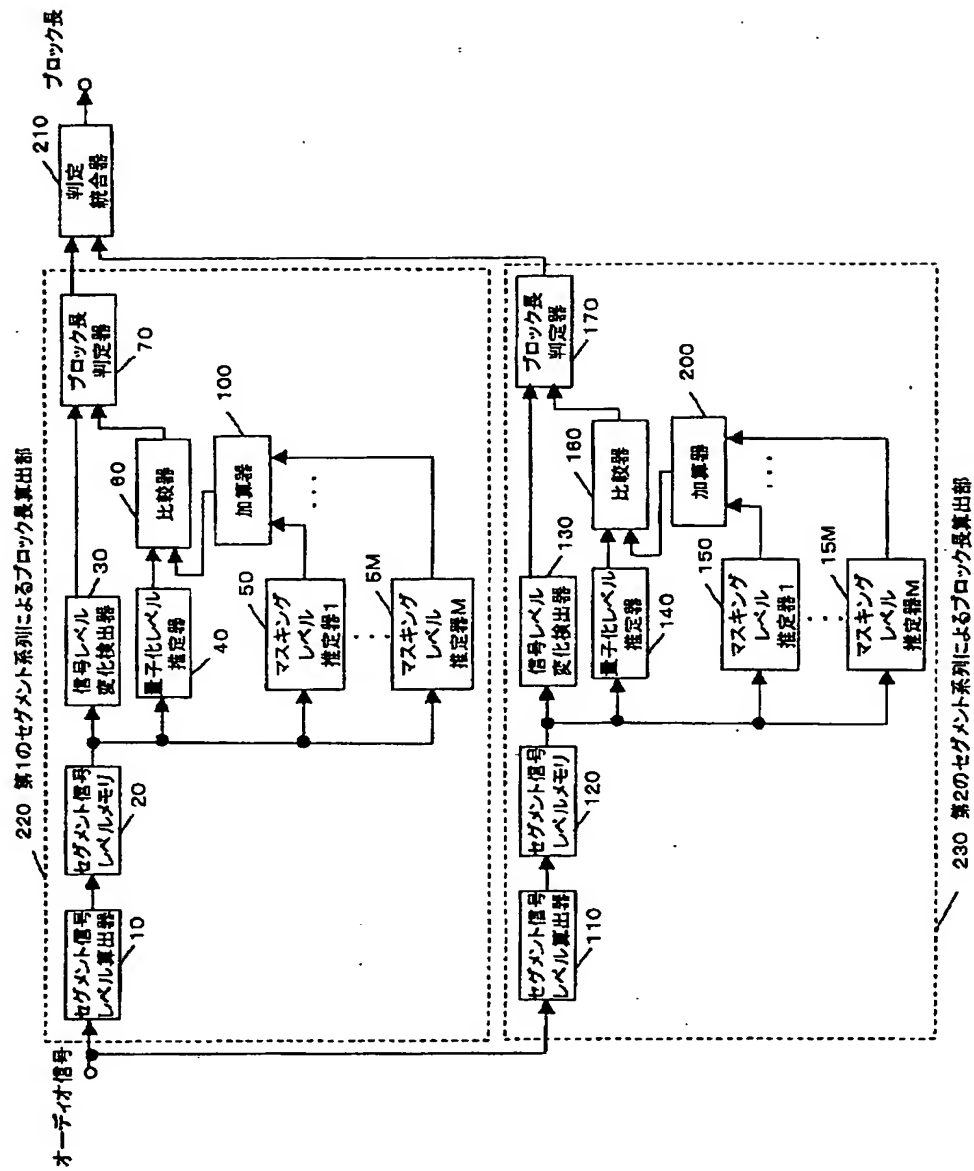
【図4】



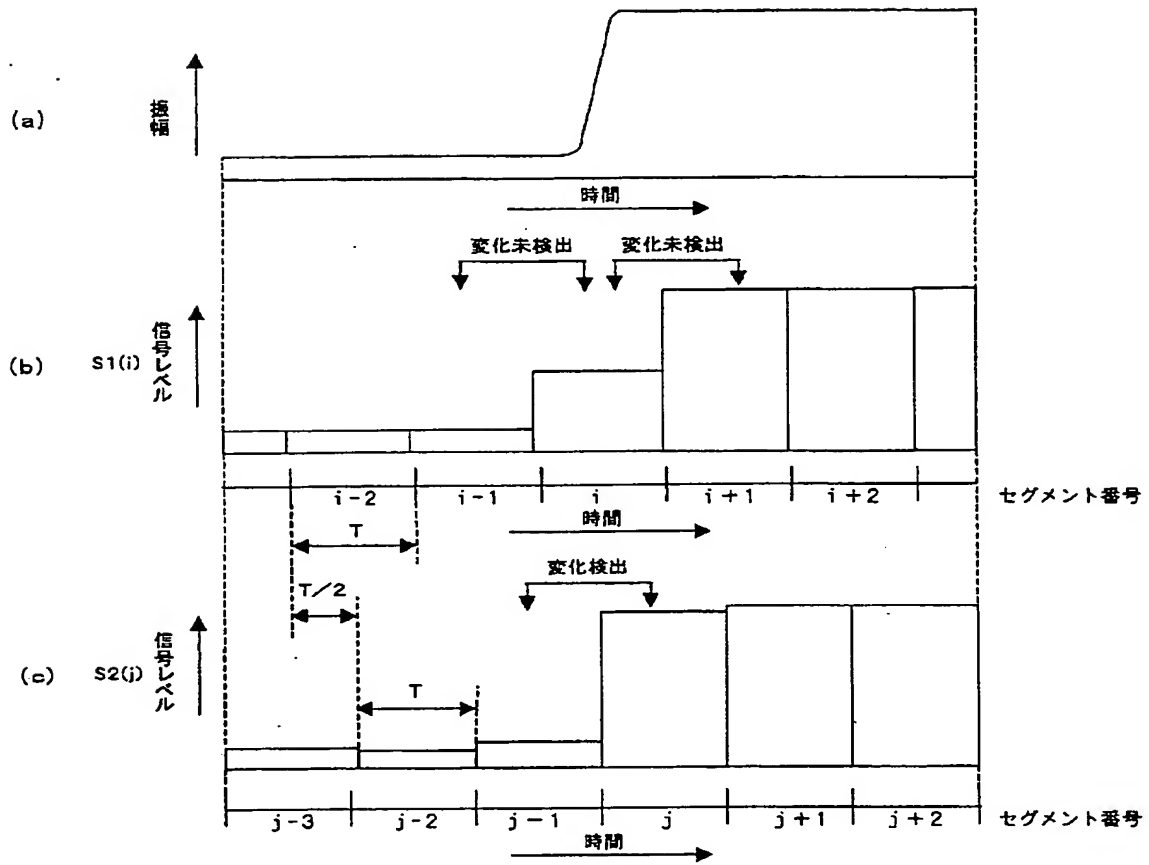
【図5】



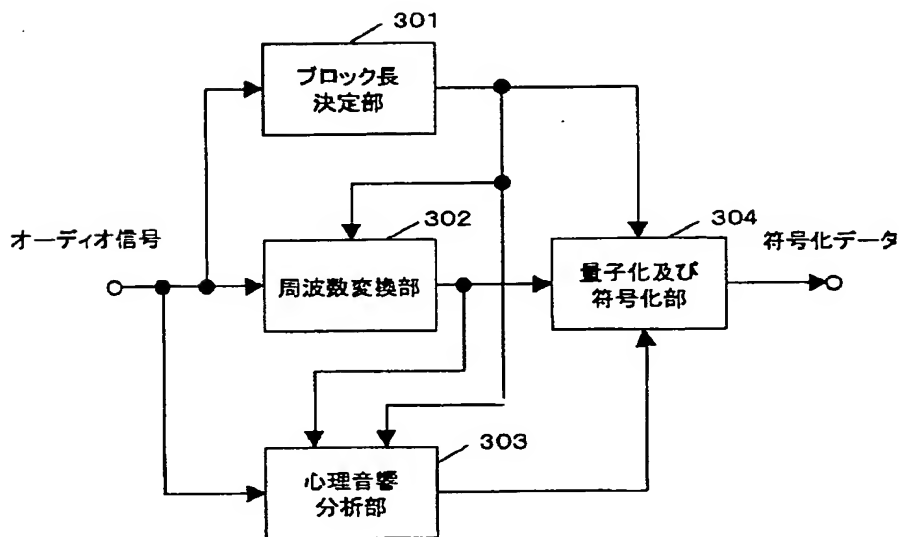
【図6】



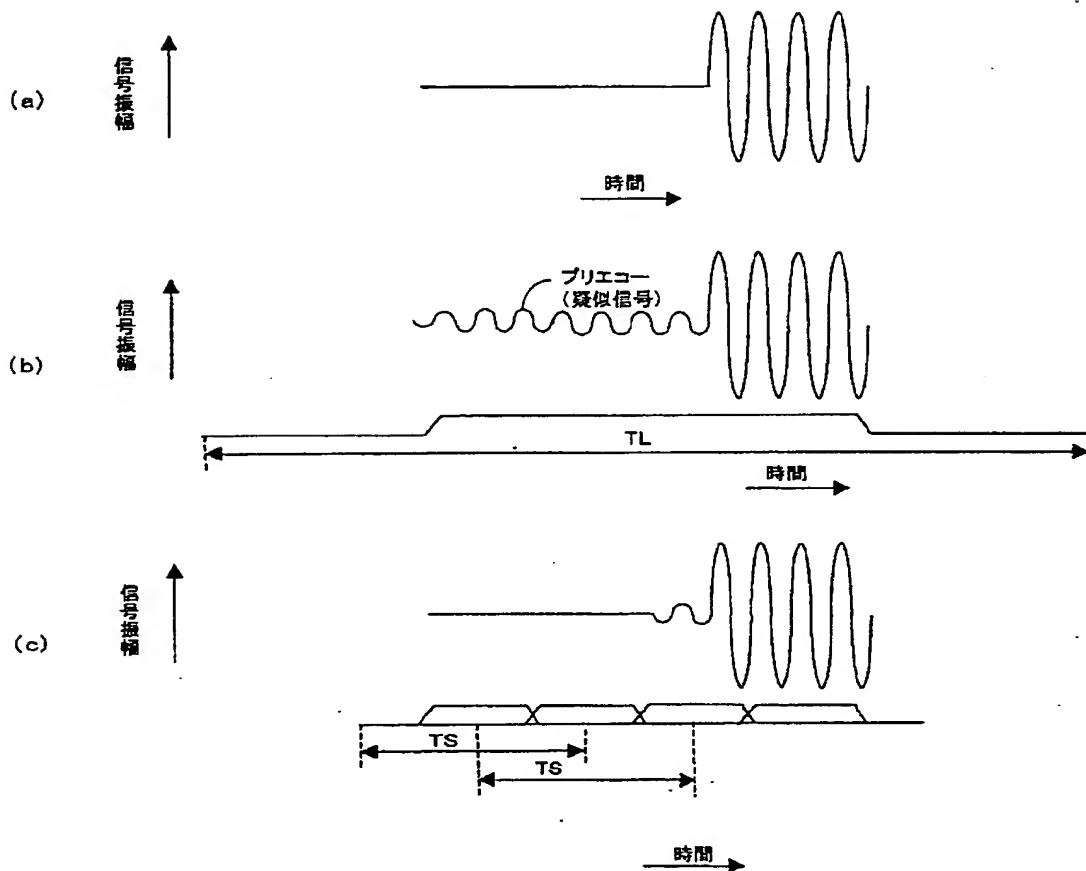
【図7】



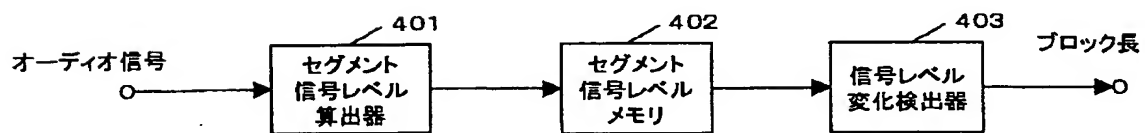
【図8】



【図9】



【図10】



フロントページの続き

(72)発明者 飯村 勝彦
大阪府門真市大字門真1006番地 松下電器
産業株式会社内

Fターム(参考) 5D045 DA20
5J064 AA01 BA16 BC08 BC11 BC14
BC16 BD02